

**ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹**

Please fill all the sections A, B and C below.

Please read Principles and Procedures Document (P & P) from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for guidelines and details before filling this form.

Please ensure you are using the latest Form from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

See also <http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest Roadmaps.

A. Administrative

1. Title:	<i>Proposal to encode Old South Arabian Script</i>	
2. Requester's name:	<i>Sultan Maktari & Kamal Mansour</i>	
3. Requester type (Member body/Liaison/Individual contribution):	<i>Liaison submission from University of California, Berkeley Script Encoding Initiative (Universal Scripts Project)</i>	
4. Submission date:	<i>28 January 2008</i>	
5. Requester's reference (if applicable):		
6. Choose one of the following:		
This is a complete proposal:	<i>Yes</i>	
(or) More information will be provided later:	<i>No</i>	

B. Technical – General

1. Choose one of the following:	
a. This proposal is for a new script (set of characters):	<i>Yes</i>
Proposed name of script:	<i>Old South Arabian</i>
b. The proposal is for addition of character(s) to an existing block:	<i>No</i>
Name of the existing block:	
2. Number of characters in proposal:	<i>32</i>
3. Proposed category (select one from below - see section 2.2 of P&P document):	
A-Contemporary ☐ B.1-Specialized (small collection) ☐ B.2-Specialized (large collection) ☐	
C-Major extinct ☒ D-Attested extinct ☐ E-Minor extinct ☐	
F-Archaic Hieroglyphic or Ideographic ☐ G-Obscure or questionable usage symbols ☐	
4. Is a repertoire including character names provided?	
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	<i>Yes</i>
b. Are the character shapes attached in a legible form suitable for review?	<i>Yes</i>
5. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?	<i>Sultan Maktari</i>
If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:	<i>Sultan Maktari, Fontlab</i>
6. References:	
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	<i>Yes</i>
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<i>Yes</i>
7. Special encoding issues:	
Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?	<i>Yes</i>

8. Additional Information:

Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see <http://www.unicode.org/Public/UNIDATA/UCD.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

¹ Form number: N3102-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?	Yes
If YES explain	<i>This version replaces WG2 N3309 (L2/07-240)</i>
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?	Yes
If YES, with whom?	<i>Alessandra Avanzini (Univ. of Pisa), Peter Stein (Friedrich-Schiller Univ. Jena), Abdullah A.A. Mikiash (Univ. Aden), Gideon Goldberg (Hebrew Univ.)</i>
If YES, available relevant documents:	
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?	No
Reference:	
4. The context of use for the proposed characters (type of use; common or rare)	rare
Reference:	
5. Are the proposed characters in current use by the user community?	yes
If YES, where? Reference:	<i>Academic publications</i>
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?	
If YES, is a rationale provided?	No
If YES, reference:	
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	Yes
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?	
If YES, is a rationale for its inclusion provided?	No
If YES, reference:	
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?	No
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?	
If YES, is a rationale for its inclusion provided?	No
If YES, reference:	
11. Does the proposal include use of combining characters and/or use of composite sequences?	No
If YES, is a rationale for such use provided?	
If YES, reference:	
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?	
If YES, reference:	
12. Does the proposal contain characters with any special properties such as control function or similar semantics?	No
If YES, describe in detail (include attachment if necessary)	
13. Does the proposal contain any Ideographic compatibility character(s)?	No
If YES, is the equivalent corresponding unified ideographic character(s) identified?	
If YES, reference:	

Proposal to encode Old South Arabian Script

Requestors: Sultan Maktari, Kamal Mansour

Under the auspices of University of California, Berkeley Script Encoding Initiative
(Universal Scripts Project) ¹

Date: 28 January 2008

Prior Proposals

In 1998, Everson submitted the first proposal for Old South Arabian (L2/98-036; N1689). An earlier version of this proposal (L2/07-240; WG2 N3309) was submitted in July 2007. Subsequently, document L2/07-287 (WG2 N3296) called into question two significant issues covered by L2/07-240: alphabetic order and character names. This proposal resolves these issues conclusively, in addition to making further refinements to the original version.

Historical Background

There is abundant evidence that Old South Arabian Script (OSA) was used before the Islamic era not only in the southwest corner of the Arabian Peninsula (modern-day Yemen), but actually in the entire Peninsula. In addition, samples of OSA have been found as far as Uruk in Mesopotamia, Delos in Greece, and Giza in Egypt. Archaeological finds show that as far back as the 8th century BCE, OSA was used in trade, religious writing, and in civil records. Following the spread of Islam in Yemen, the decline of OSA began in the 7th century CE as it was gradually supplanted by Arabic script.

OSA was typically known by the name of the then-dominant peoples in the Southern Peninsula. At various times, it was known as Sabaeen, Qatabani, or Hadramite, among others. Although it was used for a variety of languages, OSA is most strongly associated with Sabaeen. Many Peninsular languages borrowed OSA before introducing further changes of their own. Prime examples are the Thamudic, Safaitic, and Lihyanite scripts which eventually developed into independent scripts.

The westward migration of the Sabaeen people into the Horn of Africa introduced the South Arabian consonantal alphabet into the region. The transplanted script formed the roots of the Geez script of Ethiopia, which, in time and under presumably external influences, developed into a rich syllabary unlike any other Semitic script in history. Even a cursory examination of the letter forms of Modern Ethiopic writing reveal a striking similarity to South Arabian Script.

OSA inscriptions typically reveal a dominant right-to-left directionality, although there are also many cases of alternating directions, known as *boustrophedon* writing. Figure 1 is a fine example of this style of writing. OSA inscriptions were discovered early in the 19th century. Soon thereafter, two orientalist, Gesenius and Rödiger, made great strides towards deciphering the script.

¹ The authors wish to acknowledge the support of the “Universal Scripts Project” (part of the Script Encoding Initiative) funded by the National Endowment for the Humanities. The authors also wish to thank Debbie Anderson and Daniel Yacob for their generous help.

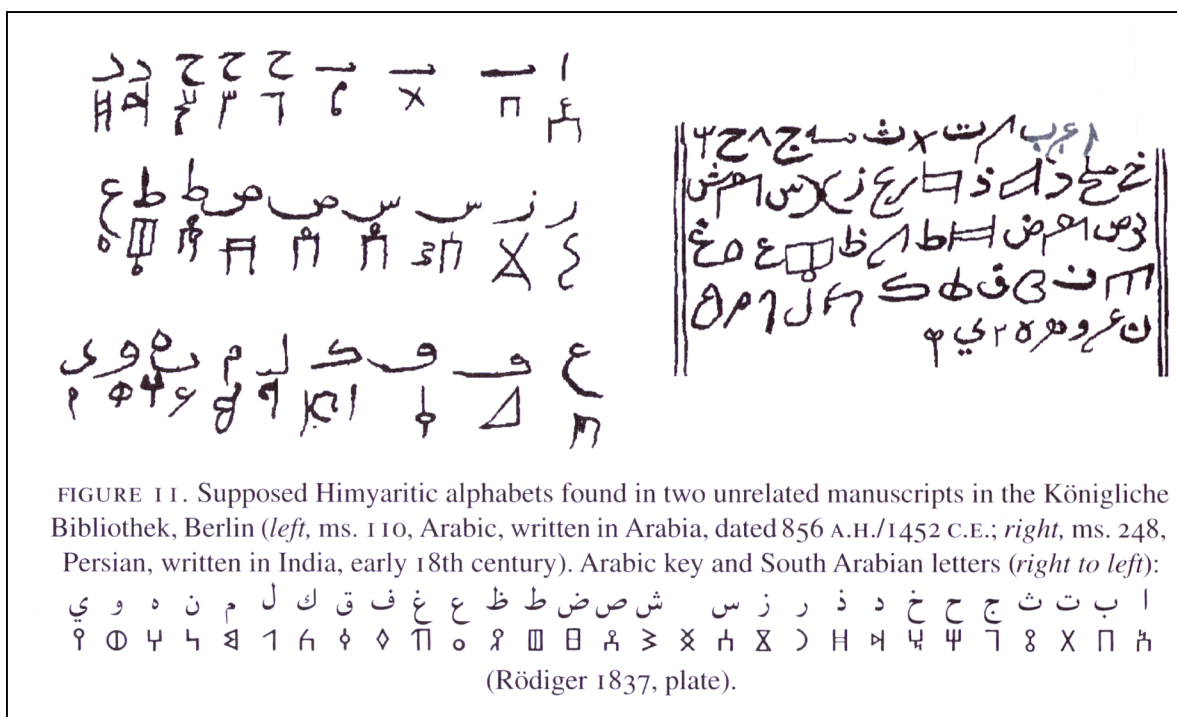
Styles of Writing

OSA inscriptions have survived primarily on stone, ceramic, and metallic surfaces. Hundreds of artifacts have been found and, to this day, continue to be discovered. Some of the best examples can be seen in figures 1–4. Figure 1—from the dust jacket of *The Cambridge Encyclopedia of the World's Ancient Languages*—shows a Sabaean inscription. Figures 2–4 depict various artifacts found in the Museum of Yemeni History in Sanaa, Yemen. Since the mid 1970s, a number of inscriptions on softer materials, such as wood and leather, have also been discovered. Although there is a significant difference between the styles of letters on the hard surfaces and those on the soft, scholars have been able to demonstrate their one-to-one equivalence. To distinguish clearly between the two styles, scholars refer to the style on hard surfaces as ‘monumental’ and the other as ‘minuscule’. The use of the term ‘minuscule’ should not be construed to mean that OSA has lower- and upper-case forms; . A better alternative to ‘minuscule’ might have been ‘cursive’. Alternatively, using terms derived from South Arabian languages, the monumental style can also be called ‘Musnad’, while the minuscule is referred to as ‘Zabur’. The following extract from Stein (2003) shows the character repertoire of OSA in a comparison of the minuscule style in the rightmost column, with the corresponding monumental style of three different eras to its left. Each line shows distinct stylistic representations of the same character.

	Monumentalschrift			Minuskel- schrift		Monumentalschrift			Minuskel- schrift
	aSab	mSab	spSab			aSab	mSab	spSab	
h	𐩣	𐩣	𐩣	𐩣	ś	𐩣	𐩣	𐩣	𐩣
l	𐩣	𐩣	𐩣	𐩣	f	𐩣	𐩣	𐩣	𐩣
h	𐩣	𐩣	𐩣	𐩣	,	𐩣	𐩣	𐩣	𐩣
m	𐩣	𐩣	𐩣	𐩣	‘	𐩣	𐩣	𐩣	𐩣
q	𐩣	𐩣	𐩣	𐩣	d	𐩣	𐩣	𐩣	𐩣
w	𐩣	𐩣	𐩣	𐩣	g	𐩣	𐩣	𐩣	𐩣
z	𐩣	𐩣	𐩣	𐩣	d	𐩣	𐩣	𐩣	𐩣
r	𐩣	𐩣	𐩣	𐩣	g	𐩣	𐩣	𐩣	𐩣
b	𐩣	𐩣	𐩣	𐩣	t	𐩣	𐩣	𐩣	𐩣
t	𐩣	𐩣	𐩣	𐩣	z	𐩣	𐩣	𐩣	𐩣
s	𐩣	𐩣	𐩣	𐩣	d	𐩣	𐩣	𐩣	𐩣
k	𐩣	𐩣	𐩣	𐩣	y	𐩣	𐩣	𐩣	𐩣
n	𐩣	𐩣	𐩣	𐩣	t	𐩣	𐩣	𐩣	𐩣
h	𐩣	𐩣	𐩣	𐩣	z	𐩣	𐩣	𐩣	𐩣
s	𐩣	𐩣	𐩣	𐩣					

Character Repertoire

The character repertoire of OSA corresponds to that of Classical Arabic, with the exception of an additional letter that is presumed analogous to the letter *Samekh* evident in West Semitic alphabets such as Hebrew and Phoenician. With striking clarity, two independent manuscripts—one dated 1452 CE from Arabia, the other 18th century CE from India—show the equivalent letters of Arabic script and Musnad. Nowadays, Semitic scholars are in agreement about the South Arabian character repertoire, though no one has yet to discover the original names of the letters. The following extract from Daniels & Bright shows the two mentioned manuscripts that helped Rödiger in deciphering South Arabian script, along with an annotation by Daniels.



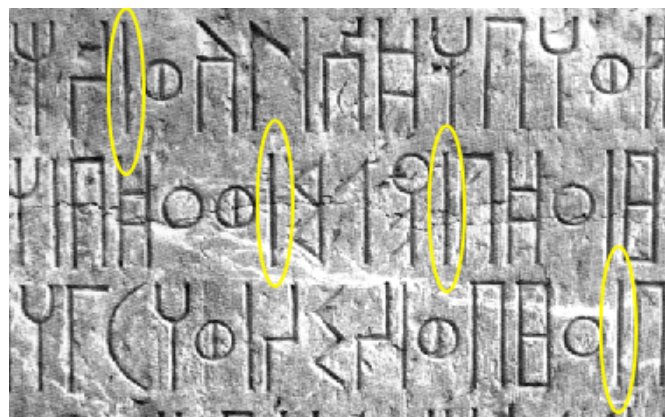
Although the historical derivation of OSA is unclear, the difference of seven letters between its repertoire and that of Phoenician makes any direct parentage between the two highly unlikely. OSA letters are all consonantal, and unlike other Semitic scripts such as Hebrew and Arabic, never developed any supplemental vocalic notation. There is no evidence of any kind of diacritic marks. It should be noted, however, that the two letters representing glides—typically transcribed as *w* and *y*—can also represent the long vowels [u] and [i], respectively. Although inscriptions have been found in varying styles, OSA—whether in monumental or minuscule form—letters are always written disjointedly and never developed any connected forms.

The following extract from Nebes & Stein shows the standard repertoire of 29 consonantal letters matching that of Daniels, but shown in South Semitic order. It should be noted that under the general category of “South Semitic”, one can find minor variations in alphabetic order.

Table 15.1 The Ancient South Arabian consonantal script

Character	Transcription	Character	Transcription
𐩦	h	𐩬	s ₃
𐩧	l	𐩭	f
𐩨	ḥ	𐩮	ʾ
𐩩	m	𐩯	ʿ
𐩪	q	𐩰	ḍ
𐩫	w	𐩱	g
𐩬	s ₂	𐩲	d
𐩭	r	𐩳	ḡ
𐩮	b	𐩴	t
𐩯	t	𐩵	z
𐩰	s ₁	𐩶	ḏ
𐩱	k	𐩷	y
𐩲	n	𐩸	ṭ
𐩳	ḥ	𐩹	ẓ
𐩴	ṣ		

While OSA inscriptions do not separate words with white space, word boundaries are clearly indicated with a vertical bar. Moreover, words can be broken arbitrarily at line boundaries without the use of any punctuation marks. In the following illustration, a few word separators are highlighted with a yellow oval. We will later demonstrate that, in a different context, this same symbol can represent the numeral 1.



Character Names

Scholars are in agreement that to this day, no one has discovered a phonemic record that establishes the names of the OSA alphabetic letters. In the previous version of this proposal (L2/07-240), we recommended the use of the Arabic names for the letters. On the other hand, document L2/07-287 proposed the use of Ethiopic instead of Arabic names since Ethiopic script is the direct descendent of OSA. However, a survey of scholarly literature does not show any use of Ethiopic names for OSA. Furthermore, all of the scholars we consulted did not consider Ethiopic names appropriate for OSA, even though some would accept the use of Arabic or North-West Semitic names. Since it is evident from the literature that scholars accept a common set of transcription characters for OSA, we propose using simple, ASCII-based Unicode character names that correspond one to one to the transcription characters and are not derived from any other language. In the case of the two characters, commonly transcribed as ' and ' , we thought it most judicious to make an exception by assigning names that are unambiguous in meaning and also common to several Semitic languages. Hence the use of *Alef* and *Ain* in the list below. The list of equivalent pairs is as follows:

h	⇒	H1
l	⇒	L
ḥ	⇒	H2
m	⇒	M
q	⇒	Q
w	⇒	W
š	⇒	S2
r	⇒	R
b	⇒	B
t	⇒	T1
s	⇒	S1
k	⇒	K
n	⇒	N
ḥ	⇒	H3
ṣ	⇒	S4
ś	⇒	S3
f	⇒	F
'	⇒	ALEF
'	⇒	AIN
ḍ	⇒	D3
g	⇒	G1
d	⇒	D1
ḡ	⇒	G2
ṭ	⇒	T3
z	⇒	Z1
ḍ	⇒	D2
y	⇒	Y
ṭ	⇒	T2
ẓ	⇒	Z2

Alphabetic Order

In L2/07-240, we had advocated the use of Arabic alphabetic order for OSA because their character repertoires of the two scripts are nearly identical. Subsequent investigation of this question has shown that scholars recognize a distinct order for OSA which differs radically from the Arabic or West Semitic order of Alef, Beth, Gimel. In fact, as a descendent of OSA, Ethiopic script gives further evidence of this order, typically known as South Semitic. Each OSA alphabetic character is correlated to its common transliteration and its proposed Unicode name. As a cross-reference only, the analogous letter in Arabic is also listed.

	Transcr Char	Unicode Name	Arabic Analog	Char
1	h	H1	Heh	ሀ
2	l	L	Lam	ለ
3	ḥ	H2	Hah	ሐ
4	m	M	Meem	መ
5	q	Q	Qaf	ቀ
6	w	W	Waw	ወ
7	š	S2	Sheen	ረ
8	r	R	Reh	ሪ
9	b	B	Beh	በ
10	t	T1	Teh	ተ
11	s	S1	Seen	ሰ
12	k	K	Kaf	ሰ
13	n	N	Noon	ነ
14	ḥ	H3	Khah	ሐ
15	ṣ	S4	Sad	ሰ
16	ś	S3	<i>Samekh</i>	ሰ
17	f	F	Feh	ቀ
18	ʾ	ALEF	Alef	አ
19	ʿ	AIN	Ain	ዐ
20	ḍ	D3	Dad	ዐ
21	g	G1	Jeem	ገ
22	d	D1	Dal	ዐ
23	ğ	G2	Ghain	ገ
24	ṭ	T3	Tah	ገ
25	z	Z1	Zain	ረ
26	ḏ	D2	Thal	ረ
27	y	Y	Yeh	የ
28	ṭ	T2	Theh	የ
29	ẓ	Z2	Zah	ሰ

The above a chart presents OSA alphabetic order as shown in Stein (2003). As previously mentioned, other sources might show slight variations. Stein’s order is based on recent work by Irvine and Beeston in which they claim “It appears that we can now claim with some confidence to have a complete South Semitic alphabet sequence running H L...”

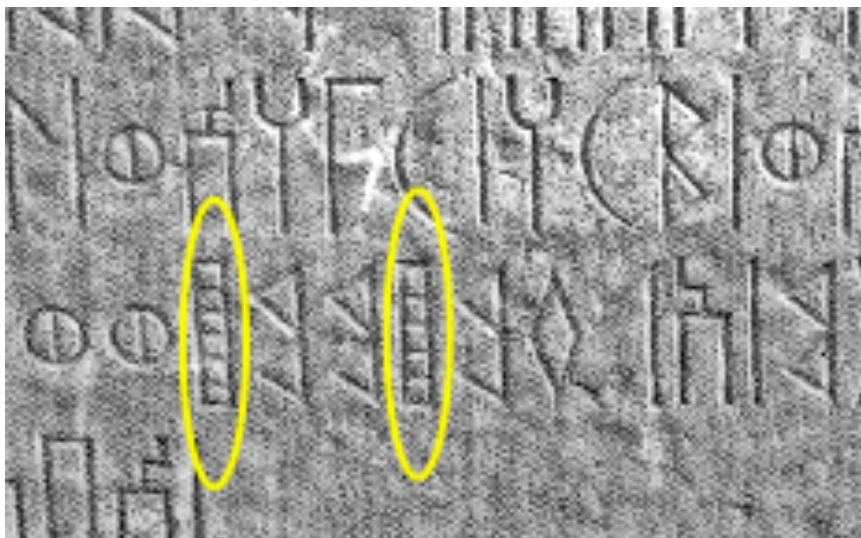
Figures

Using a set of six graphemes, all OSA numbers can be written through juxtaposition in a manner similar to Roman numerals. OSA has no symbol for zero. Following are the essential numeric graphemes alongside their meaning:

1	𐤀
5	𐤅
10	𐤆
50	𐤊
100	𐤋
1000	𐤌

With the exception of the symbols for numerals 1 and 50, the four other graphemes are used also as alphabetic characters. It is interesting to note that each of these alphabetic characters is the first letter in the name of the corresponding numeral. The symbol for 50 was evidently synthesized by removing the lower triangle in the symbol for 100.

In order to differentiate numeric quantities from surrounding text, their beginning and end are clearly marked with cross-hatched rectangular delimiters. In the following sample, the numeric indicators—highlighted with a yellow oval—surround the number 200. The surrounding text, however, indicates that the ‘200’ applies to units of 1000, making the real value 200,000 in this particular case. As evident in this inscription, the number of lines that crisscross the rectangle can vary. In other inscriptions, the lines are sometimes diagonal instead of horizontal. The numeric indicator and the word separator constitute the only form of punctuation seen in OSA writing.



Using the above-mentioned graphemes, all numeric values can be created. Unlike Roman numerals, OSA numbers show an evident right-to-left directionality. The numerals from 1 through 20 are written as follows:

1	I
2	II
3	III
4	IIII
5	𐌆
6	I𐌆
7	II𐌆
8	III𐌆
9	IIII𐌆
10	○
11	I○
12	II○
13	III○
14	IIII○
15	𐌆○
16	I𐌆○
17	II𐌆○
18	III𐌆○
19	IIII𐌆○
20	○○

How many Characters?

As we have already seen, four OSA glyphs can represent either alphabetic letters or numbers. Since the use of surrounding numeric indicators always changes their meaning from alphabetic to numeric, there can be no ambiguity in their interpretation. Consequently, there is no need to duplicate these four glyphs in the character chart for OSA.

The glyph representing numeral one resembles the one acting as word separator. In addition, the glyph can only represent the number within a clearly indicated numeric context. Since we have not found any sample inscriptions where the word separator is contrastively distinguished from the numeral one, it is appropriate to use the same character for both symbols.

The symbol for numeral fifty is unique, requiring its own character slot. As previously indicated, all other numeric values can be represented through appropriate combinations of symbols for one (ONE 𐤀), fifty (FIFTY 𐤅), five (H3 𐤆), ten (AIN 𐤁), hundred (M 𐤌), and thousand (ALEF 𐤀).

In total, 32 characters are listed in this proposal.

Future Considerations

Document L2/07-287 mentions the existence of additional “divine symbols” that appear in some OSA inscriptions. While these symbols are beyond the scope of this proposal, it would be wise to allocate a few character slots for the eventuality of their inclusion in future.

Representative Shapes

The glyph shapes used in the character chart below are meant only to indicate unambiguously each intended character, and should therefore not be considered exclusive or prescriptive in style. As we have previously hinted, the listed characters can be represented in the monumental (*Musnad*) or minuscule (*Zabur*) styles, not to mention numerous other extant variations.

Proposed Characters

1xx00	OLD SOUTH ARABIAN LETTER H1	𐩧
1xx01	OLD SOUTH ARABIAN LETTER L	𐩨
1xx02	OLD SOUTH ARABIAN LETTER H2	𐩩
1xx03	OLD SOUTH ARABIAN LETTER M	𐩪
1xx04	OLD SOUTH ARABIAN LETTER Q	𐩫
1xx05	OLD SOUTH ARABIAN LETTER W	𐩬
1xx06	OLD SOUTH ARABIAN LETTER S2	𐩭
1xx07	OLD SOUTH ARABIAN LETTER R	𐩮
1xx08	OLD SOUTH ARABIAN LETTER B	𐩯
1xx09	OLD SOUTH ARABIAN LETTER T1	𐩰
1xx0A	OLD SOUTH ARABIAN LETTER S1	𐩱
1xx0B	OLD SOUTH ARABIAN LETTER K	𐩲
1xx0C	OLD SOUTH ARABIAN LETTER N	𐩳
1xx0D	OLD SOUTH ARABIAN LETTER H3	𐩴
1xx0E	OLD SOUTH ARABIAN LETTER S4	𐩵
1xx0F	OLD SOUTH ARABIAN LETTER S3	𐩶
1xx10	OLD SOUTH ARABIAN LETTER F	𐩷
1xx11	OLD SOUTH ARABIAN LETTER ALEF	𐩸
1xx12	OLD SOUTH ARABIAN LETTER AIN	𐩹
1xx13	OLD SOUTH ARABIAN LETTER D3	𐩺
1xx14	OLD SOUTH ARABIAN LETTER G1	𐩻
1xx15	OLD SOUTH ARABIAN LETTER D1	𐩼
1xx16	OLD SOUTH ARABIAN LETTER G2	𐩽
1xx17	OLD SOUTH ARABIAN LETTER T3	𐩾
1xx18	OLD SOUTH ARABIAN LETTER Z1	𐩿
1xx19	OLD SOUTH ARABIAN LETTER D2	𐺀
1xx1A	OLD SOUTH ARABIAN LETTER Y	𐺁
1xx1B	OLD SOUTH ARABIAN LETTER T2	𐺂
1xx1C	OLD SOUTH ARABIAN LETTER Z2	𐺃
1xx1D	OLD SOUTH ARABIAN NUMBER ONE	𐺄
1xx1E	OLD SOUTH ARABIAN NUMBER FIFTY	𐺅
1xx1F	OLD SOUTH ARABIAN NUMERIC INDICATOR	𐺆

Unicode Properties of Proposed Characters

```

1xx00;OLD SOUTH ARABIAN LETTER H1;Lo;0;R;;;;;N;;;;;
1xx01;OLD SOUTH ARABIAN LETTER L;Lo;0;R;;;;;N;;;;;
1xx02;OLD SOUTH ARABIAN LETTER H2;Lo;0;R;;;;;N;;;;;
1xx03;OLD SOUTH ARABIAN LETTER M;Lo;0;R;;;;;N;;;;;
1xx04;OLD SOUTH ARABIAN LETTER Q;Lo;0;R;;;;;N;;;;;
1xx05;OLD SOUTH ARABIAN LETTER W;Lo;0;R;;;;;N;;;;;
1xx06;OLD SOUTH ARABIAN LETTER S2;Lo;0;R;;;;;N;;;;;
1xx07;OLD SOUTH ARABIAN LETTER R;Lo;0;R;;;;;N;;;;;
1xx08;OLD SOUTH ARABIAN LETTER B;Lo;0;R;;;;;N;;;;;
1xx09;OLD SOUTH ARABIAN LETTER T1;Lo;0;R;;;;;N;;;;;
1xx0A;OLD SOUTH ARABIAN LETTER S1;Lo;0;R;;;;;N;;;;;
1xx0B;OLD SOUTH ARABIAN LETTER K;Lo;0;R;;;;;N;;;;;
1xx0C;OLD SOUTH ARABIAN LETTER N;Lo;0;R;;;;;N;;;;;
1xx0D;OLD SOUTH ARABIAN LETTER H3;Lo;0;R;;;;;N;;;;;
1xx0E;OLD SOUTH ARABIAN LETTER S4;Lo;0;R;;;;;N;;;;;
1xx0F;OLD SOUTH ARABIAN LETTER S3;Lo;0;R;;;;;N;;;;;
1xx10;OLD SOUTH ARABIAN LETTER F;Lo;0;R;;;;;N;;;;;
1xx11;OLD SOUTH ARABIAN LETTER ALEF;Lo;0;R;;;;;N;;;;;
1xx12;OLD SOUTH ARABIAN LETTER AIN;Lo;0;R;;;;;N;;;;;
1xx13;OLD SOUTH ARABIAN LETTER D3;Lo;0;R;;;;;N;;;;;
1xx14;OLD SOUTH ARABIAN LETTER G1;Lo;0;R;;;;;N;;;;;
1xx15;OLD SOUTH ARABIAN LETTER D1;Lo;0;R;;;;;N;;;;;
1xx16;OLD SOUTH ARABIAN LETTER G2;Lo;0;R;;;;;N;;;;;
1xx17;OLD SOUTH ARABIAN LETTER T3;Lo;0;R;;;;;N;;;;;
1xx18;OLD SOUTH ARABIAN LETTER Z1;Lo;0;R;;;;;N;;;;;
1xx19;OLD SOUTH ARABIAN LETTER D2;Lo;0;R;;;;;N;;;;;
1xx1A;OLD SOUTH ARABIAN LETTER Y;Lo;0;R;;;;;N;;;;;
1xx1B;OLD SOUTH ARABIAN LETTER T2;Lo;0;R;;;;;N;;;;;
1xx1C;OLD SOUTH ARABIAN LETTER Z2;Lo;0;R;;;;;N;;;;;
1xx1D;OLD SOUTH ARABIAN NUMBER ONE;;No;0;R;;;1;N;;;;;
1xx1E;OLD SOUTH ARABIAN NUMBER FIFTY;No;0;R;;;50;N;;;;;
1xx1F;OLD SOUTH ARABIAN NUMERIC INDICATOR;Po;0;ON;;;;;N;;;;;

```

Note: an annotation should be added stating that NUMBER ONE (U+1xx1D) can also be used as word separator.

References

د. جواد علي. المفصل في تاريخ العرب قبل الإسلام. دار العلم للملايين. (١٠ مجلدات). بيروت. ١٩٨٧ م
Jawwad Ali. Detailed History of the Arabs before Islam. Dâr-îl-'Ilm lil-Malayîn. [10 volumes]. (in Arabic). Beirut, 1987.

عبد الله احمد عبد الله مكياش. نقوش عربية جنوبية من اليمن. رسالة دكتوراة. جامعة بغداد. ٢٠٠٢ م
Abdullah A.A. Mikiash. South Arabian Inscription from Yemen. (Doctoral Dissertation). Baghdad University, 2002.

A.F.L. Beeston. *A descriptive grammar of epigraphic South Arabian*. London, Luzac, 1962.

A.F.L. Beeston. *Four Sabaean Texts in the Istanbul Archaeological Museum*. 1952.

Florian Coulmas. *The Blackwell Encyclopedia of Writing Systems*. Blackwell Publishers. 1996.

Peter T. Daniels. “Methods of Decipherment” in Peter T. Daniels & William Bright, eds. *The World’s Writing Systems*. Oxford University Press, 1996.

A.K. Irvine & A. F. L. Beeston. “New Evidence on the Qatabanian Letter Order” in M.C.A. Macdonald & C.S. Phillips, eds. *A.F.L. Beeston at the Arabian Seminar and other papers*, Archaeopress, Oxford 2005, pp. 123-128.

Albert Jamme. *Sabaean inscriptions from Ma·hram Bilqîs (Mârib)* [with foreword by Wendell Phillips]. 19??.

Norbert Nebes & Peter Stein. “Ancient South Arabian” in Roger D. Woodard, ed. *The Cambridge Encyclopedia of the World’s Ancient Languages*. Cambridge University Press. 2004.

Peter Stein. “Untersuchungen zur Phonologie und Morphologie des Sabäischen” in *Epigraphische Forschungen auf der Arabischen Halbinsel* 3. Rahden/Westf.: Marie Leidorf, 2003.

Peter Stein. “The Ancient South Arabian Minuscule inscriptions on wood. A new genre of pre-Islamic epigraphy,” in *Jaarbericht ‘EX ORIENTE LUX’*; 39 (2005), pp. 181-199

J. Ryckmans. “Origin and evolution of South Arabian minuscule writing on wood (1).” in *Arabian archaeology and epigraphy* 12 (2001), pp. 223-235.

Wendell Philips. *Qataban and Sheba*. Victor Gollancz Ltd, London. 1955.

Museum of Yemeni History, Sanaa, Yemen

Correspondence with Alessandra Avanzini (Univ. of Pisa), Peter Stein (Friedrich-Schiller Univ. Jena), Abdullah A.A. Mikiash (Univ. Aden), Gideon Goldberg (Hebrew Univ.)

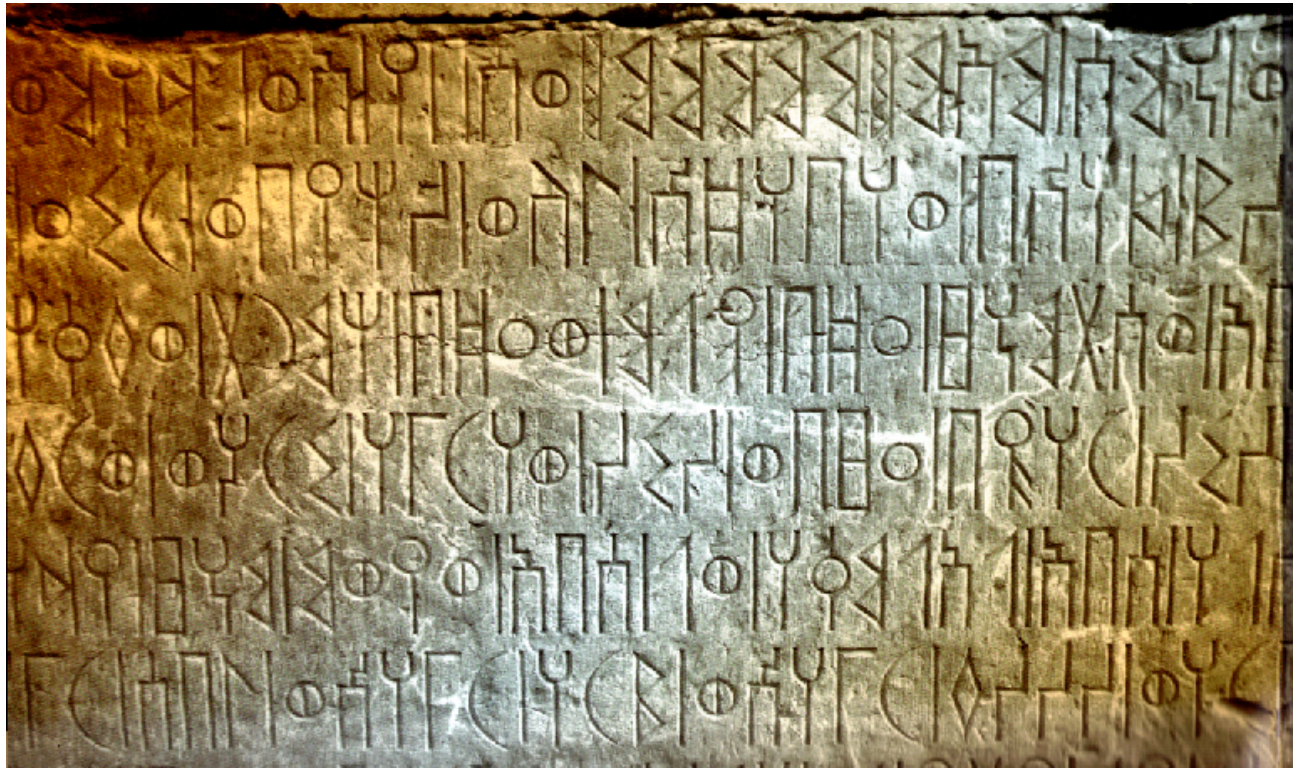


Figure 1
from the collection of the German Institute of Archaeology



Figure 2
Museum of Yemeni History, Sanaa



Figure 3
Museum of Yemeni History, Sanaa



Figure 4
Museum of Yemeni History, Sanaa

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹

Please fill all the sections A, B and C below.

Please read Principles and Procedures Document (P & P) from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for guidelines and details before filling this form.

Please ensure you are using the latest Form from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

See also <http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest Roadmaps.

A. Administrative

1. Title:	<i>Proposal to encode Old South Arabian Script</i>	
2. Requester's name:	<i>Sultan Maktari & Kamal Mansour</i>	
3. Requester type (Member body/Liaison/Individual contribution):	<i>Liaison submission from University of California, Berkeley Script Encoding Initiative (Universal Scripts Project)</i>	
4. Submission date:	<i>28 January 2008</i>	
5. Requester's reference (if applicable):		
6. Choose one of the following:		
This is a complete proposal:	<i>Yes</i>	
(or) More information will be provided later:	<i>No</i>	

B. Technical – General

1. Choose one of the following:	
a. This proposal is for a new script (set of characters):	<i>Yes</i>
Proposed name of script:	<i>Old South Arabian</i>
b. The proposal is for addition of character(s) to an existing block:	<i>No</i>
Name of the existing block:	
2. Number of characters in proposal:	<i>32</i>
3. Proposed category (select one from below - see section 2.2 of P&P document):	
A-Contemporary ☐ B.1-Specialized (small collection) ☐ B.2-Specialized (large collection) ☐	
C-Major extinct ☒ D-Attested extinct ☐ E-Minor extinct ☐	
F-Archaic Hieroglyphic or Ideographic ☐ G-Obscure or questionable usage symbols ☐	
4. Is a repertoire including character names provided?	
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	<i>Yes</i>
b. Are the character shapes attached in a legible form suitable for review?	<i>Yes</i>
5. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?	<i>Sultan Maktari</i>
If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:	<i>Sultan Maktari, Fontlab</i>
6. References:	
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	<i>Yes</i>
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<i>Yes</i>
7. Special encoding issues:	
Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?	<i>Yes</i>

8. Additional Information:

Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see <http://www.unicode.org/Public/UNIDATA/UCD.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

¹ Form number: N3102-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?	Yes
If YES explain	<i>This version replaces WG2 N3309 (L2/07-240)</i>
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?	Yes
If YES, with whom?	<i>Alessandra Avanzini (Univ. of Pisa), Peter Stein (Friedrich-Schiller Univ. Jena), Abdullah A.A. Mikiash (Univ. Aden), Gideon Goldberg (Hebrew Univ.)</i>
If YES, available relevant documents:	
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?	No
Reference:	
4. The context of use for the proposed characters (type of use; common or rare)	rare
Reference:	
5. Are the proposed characters in current use by the user community?	yes
If YES, where? Reference:	<i>Academic publications</i>
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?	
If YES, is a rationale provided?	No
If YES, reference:	
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	Yes
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?	
If YES, is a rationale for its inclusion provided?	No
If YES, reference:	
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?	No
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?	
If YES, is a rationale for its inclusion provided?	No
If YES, reference:	
11. Does the proposal include use of combining characters and/or use of composite sequences?	No
If YES, is a rationale for such use provided?	
If YES, reference:	
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?	
If YES, reference:	
12. Does the proposal contain characters with any special properties such as control function or similar semantics?	No
If YES, describe in detail (include attachment if necessary)	
13. Does the proposal contain any Ideographic compatibility character(s)?	No
If YES, is the equivalent corresponding unified ideographic character(s) identified?	
If YES, reference:	